IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): James M. Sweet et al.

Application No.: 10/608,591

Filed: 6/27/2003

Examiner: Nathan Hillery

Art Unit: 2176

Confirmation No.: 8445

Title: **DETERMINATION OF TABLE OF CONTENT LINKS FOR A HYPERLINKED DOCUMENT**

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

Sir:

DECLARATION UNDER 37 CFR §1.132

I, **James M. Sweet**, declare and state:

1. I am one of the inventors listed on the above-identified application. I reside at 394 Warren Ave, Rochester, NY, 14618.

2. I have a Bachelors and a Masters degree in Computer Engineering from Rochester Institute of Technology, Rochester, NY, USA, both received in 2003.

3. I have been employed by Xerox Corporation since 1998, where my current title is Member Research & Technology Staff II in Xerox Innovation Group. Since I joined Xerox, I have performed research and development in the area of image compression and image processing in software, and have generally specialized in software throughout both my college education and career at Xerox.

4. I have read and understand the contents of the U.S.P.T.O. Official Action of August 3, 2007, and am of the position that: a) in claims 1, 16, and 27, it would be clear to anyone skilled in the art that the "document representation" referred to in the aforementioned claims must implicitly refer to a document representation stored in computer memory; b) in claims 1, 16, and 27, it would be clear to anyone skilled in the art that the "subsequent viewing and printing" would be initiated by a user of a computer system; c) the claims being made in present application are distinctly different from and not obvious over Bharat et al. (U.S. Patent 6,112,203, hereinafter "Bharat") in view of Earl et al. (U.S. Patent 5,924,104, hereinafter "Earl"); and, d) the definition of "intra-document link" as meant in Earl, is entirely separate from the definition as meant in the present application, and that therefore Earl should not be considered in relation to the present application.

5. Claims 1, 16, and 27 each refer to "grouping the resultant set of candidate document pages into a document representation for subsequent viewing or printing of the given hyperlinked document." It would be understood by anyone skilled in the art that any meaningful representation of a hyperlinked document would necessarily have to be stored in memory, due to the inherent nature of hyperlinking. Merrian-Webster defines a "hyperlink" as "an electronic link providing direct access from one distinctively marked place in a hypertext or hypermedia document to another in the same or a different document." It would be fundamentally impossible for any document not stored in memory to contain a hyperlink, since by definition a hyperlink is an electronic link providing direct access to another document or place in the same document. This type of electronic direct access is impossible in any document not stored in memory. Therefore, because the "document representation" is said in the claim to represent a "hyperlinked document," it is implicit in the claim that the document representation must be stored in memory.

6. Furthermore, the "subsequent viewing or printing of the given hyperlinked document" would be reasonably understood by anyone skilled in the art to have been initiated by the user of a computer system. While it would be theoretically possible for a human to manually translate the coded representation of the hyperlinked document into a viewable form, this is impractical to the point of absurdity. It would be understood by anyone skilled in the art that "viewing or printing" a "hyperlinked document" would be an activity performed by a computer and initiated by a user of that computer.

7. It is my understanding that the Examiner believes the "intra-document links" referenced in several of the claims are obvious from Bharat, because both inventions take into account whether two nodes reside on the same server when determining to retain a link between those nodes. I believe this interpretation to be incorrect for a number of reasons. It should first be noted that the links being retained in Bharat have a very different meaning than the links being retained in the present invention. In Bharat, the intention is to improve the ability of a search engine to retrieve pages which are "important" in the context of the subject matter. Bharat has wisely observed that one's opinion of one's own importance is perhaps less accurate than others' opinions of one's importance, and has thence chosen to completely discard links that reside within the same server. Those skilled in the art will immediately recognize this technique, similar to the PageRank algorithm used by Google, as a means to prevent a common search engine exploit, wherein a website artificially increases its own perceived importance by linking to itself many times. In contrast, the links being retained in the present invention are meant to indicate a very specific and close relationship between hyperlinked pages, i.e. that they are part of the same continuous document. Bharat seeks to retain only those links which do not have a close relationship, while the present invention seeks to retain only those links which do have a specific type of close relationship.

8. Much discussion has been made of the use the phrase "weed out" in the present application, used on page 7, lines 9-22. It is my understanding that the Examiner does not believe that the language "weed out" implies removal of the links in question, asserting instead that "weed out" only implies identification, not removal. The full claim language phrase in question is "weed out links which have properties that are not characteristic of intra-document links." It should be noted that the following claim language makes clear that the intention of this step is to produce "a list of intra-document links." It defies the imagination to come up with any reasonable interpretation that, seeking to produce a "list of X," would first go to great pains to identify items which "have properties that are not characteristic of X," and yet then proceed to just carelessly toss those same items into the final "list of X" anyway, disregarding the previous identification.

9. I would also point to the present application, page 10, lines 23-27, where the applicants refer to the identification of hyperlinks that "are significantly different in a property that is typical of intra-document links," going on to again refer to the "link to a page on a different server" as being one of these criteria. In this part of the specification, the applicants used the phrase "filtered out," as an alternative to the "weed-out" language which will also further reinforce for those skilled in the art as to what the Applicants mean. Regardless, it would be clear to any skilled in the art that the Applicants are seeking to remove links that do not fit the characteristics of intra-document links.

10. Having established that Bharat uses the criteria of co-residence on the same web server as primarily an exclusionary criteria, while the present invention uses it as primarily an inclusionary criteria, one not skilled in the art may at this point be tempted to say that the links being retained in Bharat are merely the inverse of the links being retained in the second invention. This perception would be false. Recognizing that a close relationship may artificially inflate importance, Bharat seeks to eliminate *all* types of close relationships, and therefore, the inverse would be to include all types of close relationships indiscriminately. The present invention seeks to identify a *specific type* of

-4-

close relationship (the relationship of being part of the same document) and therefore the inverse would include a number of links which also identified a close relationship of a different type, i.e. a different set of links than that which is taught by Bharat.

11. One skilled in the art would further recognize that whether or not two given hypertext pages reside on the same web server is merely a piece of data, and not a teaching in and of itself. To suggest that Bharat makes obvious the present invention because both sample this piece of data is analogous to suggesting that a design for a central air conditioning system makes obvious a design for a nuclear reactor, because both inventions require the measurement of temperature in order to maintain proper operation.

12. In short, even if the Applicants' definition of "weed out," is or is not accepted, it is nevertheless irrelevant to whether Bharat makes the present invention obvious. Making a determination based on whether two pages reside on the same web server, is not original to the teachings of Bharat, nor the present Applicants, as it is a piece of data that is used everyday in a number of different applications, such as Domain Name Server lookups, corporate firewalls, page caching mechanisms, web browser pop-up blockers, etc. It will be clear to those skilled in the art that many unique inventions will at times sample some of the same data. It is the different <u>ways</u> in which this and other data points are interpreted that is central to both Bharat and the present invention and which is also where the distinction between them is to be found.

13. It is unfortunate that both the present Applicants and Earl have chosen to use the term "intra-document link." It will be clear to anyone skilled in the art that this is merely a semantic coincidence, stemming from a different meaning of the word "document."

14. The present Applicants define an intra-document link as "links within a given web page that may link to *other* pages within the same document" (emphasis added).

This makes it clear that a "document" as intended by the current applicants may refer to a group of more than one web page, since it would be impossible to link to other pages within a document if a document only ever comprised a single page.

15. Earl does not ever specifically define "document" or "intra-document link," however, one skilled in the art can clearly discern what is meant by "document" in Earl's teachings. In Earl, Column 3, Line 44 through Column 4, Line 15, five examples are given of links which may be "intra-document." Earl observes that only the first two can reliably be said to be "intra-document," while all five could be "intra-document." One skilled in the art will readily observe the point Earl is making: In all five examples, the link appears to be a link to the same web page as the one in which the link appears. However, due to the details of how link destinations are interpreted by web servers, only the first two examples are certain to link to the same web page as the one in which the link appears. It is abundantly clear from this that what Earl means by "intra-document" is a link whose destination is the same as its source, i.e. a circular link to the same page (hereinafter referred to as "circular link").

16. One skilled in the art will recognize that Claim 1 of the present application both implicitly and explicitly excludes circular links from any form of consideration. Claim 1 teaches of "a recursive application of page-level link analysis to the linked candidate document page and any *further* nested candidate document pages thereby identified" (emphasis added). The word "further" explicitly excludes circular links from consideration, because they would not increase the number of candidate document pages identified. In addition, the "recursive application" will be recognized by one skilled in the art to implicitly exclude circular links, since the inclusion of circular links in such an analysis would result in infinite recursion.

I, the undersigned, further declare that all statements made herein are of my own knowledge are true and that all statements made on information and beliefs are believed to be true; and further, that these statements were made with the knowledge that willful

false statements and the like so made are punishable by fine or imprisonment, or both under Section 1001 of Title 18 of the United States Code, and further such willful statements may jeopardize the validity of the application or any patent issuing thereon.

Respectfully submitted,

James M. Sweet

11/13/07

Date